DOCUMENT RESUME

ED 160 734 CE 015 199

AUTHOR Misanchuk, Earl R.
TITLE Uses and Abuses of Evaluation in Continuing Education Programs: On the Frequent Futility of Formative, Summative, and Justificative Evaluation.
PUB DATE Apr 78
NOTE 22p.; Paper presented at the Adult Education Research Conference (San Antonio, Texas, April, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Adult Education; *Adult Education Programs; Evaluation Methods; Evaluation Needs; *Formative Evaluation; *Guidelines; *Program Evaluation; *Summative Evaluation
IDENTIFIERS *Confirmative Evaluation; *Justificative Evaluation; Scriven (Michael)

ABSTRACT

Inappropriate uses in continuing education of Scriven's (1967) concepts of formative and summative evaluation are described by the author, and guidelines are presented to suggest when formative and summative evaluations are appropriate in continuing education programs. After clarifying what is meant by formative and summative evaluation, the author describes and explains the role of confirmative evaluation, an evaluation analagous to formative and summative evaluation but pertaining to programs that have been implemented for some time and are now up for review. A non-valid evaluation role, justificative evaluation, is also described and explained, and guidelines for obviating a need for that role are suggested. (Justificative evaluation is defined as an attempt to justify the design after the fact.) Finally, a list of conditional rules suggesting when to and not to use formative, summative, confirmative, and justificative evaluations is presented. (Author/BM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED160734

USES AND ABUSES OF EVAULATION IN CONTINUING EDUCATION PROGRAMS: ON THE FREQUENT FUTILITY OF FORMATIVE, SUMMATIVE, and JUSTIFICATIVE EVALUATION

Earl R. Misanchuk
Extension Division
The University of Saskatchewan

Paper presented at the Adult Education Research Conference
San Antonio, Texas, April, 1978
© E.R. Misanchuk, 1978

OE 015199

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Earl R. Misanchuk

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ABSTRACT

Inappropriate uses in continuing education of Scriven's (1967) concepts of formative and summative evaluation are described and guidelines are presented for when formative and summative evaluations are and are not appropriate in continuing education programs. A role of evaluation analagous to formative and summative evaluation but pertaining to programs that have been implemented for some time and are now up for review, dubbed confirmative evaluation, is described and explained. A non-valid evaluation role, justificative evaluation, is also described and explained, and guidelines for obviating a need for that role are suggested.

To comprehend the nature and origin of current abuses of evaluation in continuing education it is helpful to review the phenomenal progress of that newest growth industry in education, be it continuing or preparatory—educational evaluation. A decade ago or so, in most circles the word evaluation meant the systematic gathering of evidence that a student had learned (or failed to learn) what he or she was expected to learn. In other words, evaluation dealt with pupil achievement.

During the middle and late sixties, some rethinking of the role and scope of evaluation was done by some noteworthy scholars (e.g., Cronbach, 1964; Scriven, 1967; Stake, 1967) and the focus of evaluation was subsequently expanded to include the instructional materials, curriculum, teaching strategies, and other variables in a teaching/learning system. Cronbach's (1964) suggestion that the components of a developing curriculum become a focus for evaluation data, and Scriven's (1967) expansion of that thought into the distinction between formative and summative evaluation represented a quantum leap in evaluation thinking. As aspects of the teaching/learning system other than student achievement were proposed as being eligible for evaluation, it became commonplace to speak of the new, almost separate, fields of program evaluation and product evaluation.

At first, seniority played as large a part in the evolution of educational evaluation as anywhere else, and among the first (and most-widely respected) of the program evaluators were those who had a firm grounding in the evaluation of student achievement.

Student achievement, being one of the outcomes of a program, must indeed reflect the quality of that program (the argument went) and must perforce be included in any evaluation of the program; evaluators of student achievement saw the movement as simply a widening of their roles. Besides, who better qualified to collect quantitative data and analyze them than those who had already been doing it for somewhat different reasons?

The broadening of the focus of educational evaluation, however, had the effect of proliferating evaluation models, among them those that eschewed quantitative methods in favor of others (e.g., Eisner, 1972, 1975, Note 1; McCutcheon, 1976; Rippey, 1973; Schwille & Porter, 1976; Vallance, 1976; Wolf, Potter & Baxter, 1976). Indeed the yardstick used for measurement, it was suggested, might sometimes be ignored deliberately (Scriven, 1973; Stake, Note 2). The creation of these new models has changed the complexion of educational evaluation considerably in the last decade, but one of the few things that has remained relatively constant while the field has been evolving is the raison d'être of evaluation: most evaluators agree that the function of evaluation is to permit informed decision-making.

The movement of educational evaluation gathered momentum, and figured so prominently in the then current hot issue of accountability in decision-making, that agencies other than educational ones looked to the models developed in educational evaluation for salvation with respect to their own concerns about

accountability. Consequently, social action programs of various types adopted much of the terminology and the orientation of education evaluation, often, apparently, without carefully examining the consequences of the shift of focus from educational curricula to social action programs. Models developed by Stake (1967) and Stufflebeam (1971) were sometimes adopted wholesale. Sometimes the fit was relatively good; at other times imaginations had to be stretched to adapt the reality to the model. As products and programs to be evaluated grew more divergent in nature from the original basis on which most evaluation theory was originally forged--school curricula--the sense of the concepts and terms used in various evaluation models was distorted by greater and greater amounts. One would hear of attempts to do "summative" evaluations of unique (in the sense that they would not happen again) events.

Whether deliberate or not, this distortion has unfortunate consequences for communication within the field. For evaluators and program planners to be able to accurately discuss evaluation in their projects requires a common terminology. If you will excuse a mashed metaphor, a rose by any other name may smell as sweet, but the term 'formative evaluation' implies certain attendant conditions which have not always been met in recent practice.

Prerequisites for Formative Evaluation

By definition, formative evaluation is evaluation whose role is to "...discover deficiencies and successes in the intermediate versions of a new [product]...(Scriven, 1967, p. 51)"; formative evaluation is a part of the process of development of a new product, be it a curriculum, a textbook, a television series, or a seminar. I want to emphasize two things about the definition: the first is that it takes place during the development of the product; the second is that it takes place with respect to a product. These two points sometimes get overlooked when we speak of formative evaluation in continuing education contexts.

Let's consider the first point--that formative evaluation takes place during the development of the product. The systematic development process (e.g., see Baker, 1974, 1977; Baker & Schutz, 1971; Markle, 1967) requires that an instructional product be conceived and largely developed in isolation from the target learner: a first draft is devised, making it as good as possible by taking it through as many revisions as are judged necessary to bring it to the prototype stage; the prototype is then given, under careful supervision, to one or two learners (who are not members of the target audience themselves, but who are similar to those members) to subject it to a preliminary field test. All of this happens <u>before</u> the design is implemented with the learners themselves--to do otherwise would raise serious moral and ethical concerns.

(To be sure, it may be--ought to be--field tested later with a larger number of learners, probably sampled from the target group, but by this time the design ought to be fairly stable; few modifications to the basic design are to be expected at this stage.)

Now let's consider the second point. Note that formative evaluation is meaningful only when the object of evaluation is a product--a durable, replicable entity. Typically, we think of a product as a tangible object (e.g., a book, a film, a handout, etc.). However, it is also possible to conceive of a sequence of activities or a series of steps (i.e., what we frequently call a process) as a product, so long as the sequence of activities or series of steps can be repeated time and again with only minor, if any, variations. This interpretation, which will be used throughout this paper, permits the consideration of a person's activities as a product, so long as the condition of relative invariance is met.

The product, whether tangible or not, must have, above all, predictable and replicable effects (i.e., it must be consistent in its operation) if we are to apply formative evaluation to it. Simply put, formative evaluation can be conducted on an educational product if and only if the product can be applied time and time again in substantially the same form and manner (i.e., it is durable) with substantially the same result (i.e.,

its effects are replicable). This means that it is not meaningful to speak of formative evaluation with respect to an activity that is interactive or dependent (e.g., as is the case when instruction takes different directions in response to certain events, such as expressions of learner interest). I am not saying that teaching should never be interactive; I am saying that if it is, it is meaningless to speak of its formative evaluation.

Summative and Confirmative Evaluations

Summative evaluation, on the other hand, is the evaluation of a finished product (Scriven, 1967, p. 40-43). The product, by the time summative evaluation is performed, has been developed and field tested, and is in more or less immutable form. The basic decision to be made on the basis of a summative evaluation is "Should we implement this product (or program, seminar series, etc.)?" A summative evaluation demonstrates for public consumption what the product is capable of doing. The information will not be used by the product developers to make modifications to the product; it will be used by administrators to decide whether or not they will adopt the product and put it into use.

There is another evaluation role, one extant in both preparatory education and in continuing education but perhaps more evident in the latter, that Scriven's distinction does not cover--the situation in which a product (which has undergone formative,

and possibly summative, evaluation) has been implemented and operated for some time, and the time has come to review its efficacy. The evaluation question to be asked is "Should this program be discontinued, and, if so, should it be replaced?" The negative answer to the first part of the evaluation question begs the second question: "If it should be continued, should it be continued in its present form, or should it be modified (and if so, how)?" This role of evaluation, which I propose to call confirmative evaluation, is substantially different from either formative or summative evaluation. There is, first, the aspect to confirmative evaluation of final judgment (accept it as is, or reject it) that is implicit in the question "Should the program be discontinued?", as there is in a summative evaluation. But there is also an aspect of formative evaluation implicit in a confirmative evaluation if the answer to the first question is negative, since the second question asks how the program might be revised, if indeed revision is indicated. This contingency of evaluation questions makes the role of confirmative evaluation different from the roles of either formative or summative evaluation.

To call a confirmative evaluation a summative evaluation followed by a formative evaluation seems inconsistent with the original meaning of the words: by definition, formative evaluation takes place before the completion of a product, while summative evaluation takes place after completion. Ergo, a formative

evaluation of a product cannot succeed its summative evaluation.

Although it is possible to conceive of a situation in which summative evaluation is done without any formative evaluation having preceeded it, such a state of affairs seems quite inconsistent with the situation Scriven described. Similarly, it would be possible to conceive of a confirmative evaluation being done on a program that had the benefit of neither formative nor summative evaluation. In such a situation, it might be tempting to regard the confirmative evaluation as formative. However, in the reality of program development, formative evaluation is seldom, if ever, conducted with the intention (or even possibility) of discontinuing the development of the program in response to evaluation findings. In other words, the program development process is such that, if properly executed (i.e., if a proper needs identification is done initially), there is very little likelihood that formative evaluation information would lead to the discontinuance of program development. Rather, formative evaluation findings would lead to modification of the program being developed. In confirmative evaluation, of course, a very real possibility exists that the data would indicate that the program should be discontinued, without replacement. Hence the distinction among formative, summative, and confirmative evaluations seems necessary.

In addition, there are pragmatic concerns that help to

differentiate confirmative from formative and summative evaluation. Who should perform confirmative evaluations--someone intimately involved with the program (as in formative evaluation) or someone quite divorced from it (as in summative evaluation)? It would seem that the relatively high degree of objectivity presumed of the summative evaluation ought to be applied to the confirmative evaluation decision regarding whether the on-going program ought to be terminated or continued, but the formative evaluator would be in a much better position to make the judgment whether the program ought to continue intact or be modified (and how, if it is), given that the first part of the evaluation question indicated continuance.

At the present time, there are no definitive answers to the question. Hopefully, further experience with confirmative evaluations will provide the needed guidelines.

Justificative Evaluation

Finally, I have encountered on numerous occasions a demand for an evaluation role which I believe ought not to be filled at all.

In continuing education programs, it is not uncommon to find a type of evaluation in which questions such as the following are asked: "How relevant did you find the section on. . .?"; "How useful will the information learned in this workshop be in your present job role?"; or, even more blatantly, "Do you feel you

learned enough during this seminar to justify your attendance?" Aside from pratically begging the question, such inquiries, typically made at the conclusion of the learning experience, actually ask for information that should have been gathered prior to the design of the experience. I have dubbed such "evaluation", attempts justificative evaluation, because they seem most often to be attempting to justify the design of the event after the fact.

Despite the name I am suggesting, I do not believe that justificative evaluation should be considered a valid evaluation role at all. The identification of needs is an important part of the planning process, but the need identification process ought to happen before, not after, the design and implementation of the event.

Furthermore, it is usually possible to get <u>some</u> positive response about an educational experience at its conclusion (if nothing else, participants are rather glad it's over!), but it seems to stretch credibility that this data could be considered part of objective needs identification. Most people typically confuse pleasant experiences with valuable ones and are therefore likely to judge the worth of the learning in terms of its enjoyment.

Returning to the stated purpose of evaluation--to make decisions--and examining justificative evaluation in the light of that purpose, I can only conclude that justificative

evaluations are purposeless. Any decisions about the design of the experience that come out of the process are bound to be useless with respect to the experience itself, since the experience is already over. No knowledge of value vis à vis the needs of the learners has been gained that could not have been gained more easily and more accurately by other means (i.e., by doing an adequate needs identification). (Those who are tempted to argue that something is learned from the process--the efficacy with which the programmer can "sense" what needs exist--are confusing evaluation of the program with the evaluation of the programmers' ability to identify needs mystically.)

The close relationship between the identification of educational needs and the process of evaluation has been noted (Kaufman, 1977). However, the two processes are not identities; rather, they appear to be two sides of the same coin. They share some important similarities--their general raison d'être of providing information for decision-making, and their requirements for objectivity, validity, reliability, and utility--but they also have an important difference: needs identification ought to take place prior to the planning phase, whereas evaluation ought to take place during the development phase (formative evaluation), immediately after the development phase and before the implementation phase (summative evaluation), or during the implementation phase (confirmative evaluation). Thus, the need for justificative evaluation is obviated.

Should Evaluation Be Done?

Weiss (1972) identifies several types of uses for evaluation, couched in terms of particular decisions to be made, that are well worth reiterating:

1. To continue or discontinue the program
2. To improve its practices and procedures
3. To add or drop specific program strategies and techniques
4. To institute similar programs elsewhere
5. To allocate resources among competing programs
6. To accept or reject a program approach or theory (p. 16-17)

Couched in terms of the three role descriptors discussed earlier, Weiss' list, essentially a list of implicit questions which constitute reasons for doing evaluations, can be classified as formative (2, 3, 6), summative (4, and possibly 5 depending on the context), or confirmative (1, possibly 2 if it immediately follows 1 in an on-going program, and possibly 5, depending on the context) evaluation questions. The list suggests reasons why an evaluation might be done, but does not suggest any conditions under which it might be reasonable to consider doing the evaluation. Such a list of conditions would be useful. In the past half-dozen years, evaluation has become a buzz-word in public agencies of various kinds. For some reason (lack of clarity in

communication?, lack of understanding of the concepts?) evaluation requirements are often attached willy-nilly to projects, with seemingly little concern for their usefulness or meaning. Projects that should not or cannot be evaluated meaningfully are being required to have evaluation components. Perhaps the following list of suggested conditional rules will help to ameliorate the problem:

1. Formative evaluation should be performed only if the object of evaluation is a durable, replicable entity. If the entity is not capable of being presented time and again in largely the same form and manner with largely the same effect, formative evaluation is probably meaningless and a waste of time and money. Although durable, replicable entities are often referred to as products, recognize that human beings and the processes they engage in can well be part of the package, provided that their actions are replicable and that their involvement in the activity is durable.
2. Summative evaluation should also be performed only if the object of evaluation is a durable, replicable entity that has been subjected to formative evaluation earlier. Summative evaluation is designed to demonstrate what a product is capable of doing in extant form. Obviously if the form is malleable (i.e., the product is

not durable and/or replicable) summative evaluation is a waste of time since there is no reason to believe that subsequent or previous applications replicate the application under evaluation. Furthermore, to perform a summative evaluation an a (durable, replicable) product that has not been subjected to formative evaluation is relatively cost-ineffective. Negative results from such a summative evaluation could be attributed to sub-optimal design that could perhaps be remedied easily and inexpensively if the flaw were determined through formative evaluation.

3. Confirmative evaluation may be performed on a product or program that has not had the benefit of either formative or summative evaluation (although the fact that it is possible should not be taken to mean that it is desirable to implement unevaluated products or programs and later perform confirmative evaluation on them). But confirmative evaluation should be undertaken only on a program that has been in place for some time and whose current efficacy is validly in doubt. Furthermore, the program should have the capability of operating unchanged for another period of time. Obviously, the periods of time and validity of doubt can only be judged in relation to such variables

as cost, importance of the program content, number of people involved, etc.

4. Justificative evaluation should not be performed under most circumstances. Prior identification and analysis of learning needs is far more preferable and ethically responsible.

Implicit in the above set of rules are the following, almost heretical notions:

1. Neither formative evaluation nor summative evaluation ought to be attempted for a unique ("one-shot") event. Since a "one-shot" event cannot be subjected to confirmative evaluation, and justificative evaluation should never be undertaken, it follows that "one-shot" events ought not to be evaluated at all.
2. Formative evaluation should not be undertaken unless there is a high probability that the object of evaluation (event, product, etc.) will be repeated in the immediate future. Without the real prospect of durability and replicability of effects, formative evaluation is a waste of time and money.
3. Summative evaluation should be undertaken only when you are considering adopting a product developed elsewhere, or are considering marketing a product that you have developed. In the latter case, the

evaluator should be external to the developing agency or group (i.e., you should not do the summative evaluation of the product you developed).

If the suggestions in the paragraph above were taken seriously and applied widely in continuing education, we would probably have far fewer evaluations taking place, but the proportion of useful and meaningful evaluations would probably rise dramatically.

REFERENCE NOTES

1. Eisner, E.W. (n.d.) Educational connoisseurship and educational criticism: Their form and functions in educational evaluation. Stanford University: Unpublished paper.

2. Stake, R.E. (1972) Responsive evaluation. Unpublished paper, December 5, 1972.

REFERENCES

Baker, E.L. (1974) Formative evaluation of instruction. In W.J. Popham (Ed.), Evaluation in education: Current applications. Berkeley, Calif.: McCutchan.

Baker, E.L. (1977) The dance of evaluation: Hustle or minuet? Journal of Instructional Development, 1(1), 26-28.

Baker, R.L., & Schutz, R.E. (Eds.) (1971) Instructional product development. New York: Van Nostrand Reinhold.

Cronbach, L.J. (1964) Course improvement through evaluation. Teachers College Record, 64, 672-683.

Eisner, E.W. (1972) Emerging models for educational evaluation. The School Review, 80, 573-589.

Eisner, E.W. (1975) The perceptive eye: Toward the reformation of educational evaluation. Invited Address, Annual Meeting of the American Educational Research Association, Washington, March, 1975.

Kaufman, R. (1977) Needs assessments: Internal and external. Journal of Instructional Development, 1(1), 5-8.

Markle, S.M. (1967) Empirical testing of programs. In P.C. Lange (Ed.), Programed instruction: The 66th yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.

McCutcheon, G. (1976) The use of ethnography and criticism as methods for disclosing classroom settings. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April, 1976.

Rippey, R.M. (Ed.) (1973) Studies in transactional evaluation. Berkeley, Calif.: McCutchan.

Schwille, J., and Porter, A. (1976) The use of fieldwork in educational evaluation: From the perspective of two nonanthropologists. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April, 1976.

Scriven, M. (1973) Goal-free evaluation. In E.R. House (Ed.), School evaluation: The politics and process. Berkeley, Calif.: McCutchan.

Scriven, M. (1967) The methodology of evaluation. In R.W. Tyler (Ed.), Perspectives of curriculum evaluation. Chicago: Rand McNally.

Stake, R.E. (1967) The countenance of educational evaluation. Teachers College Record, 68, 523-540.

Stufflebeam, D.L. (Chmn.) (1971) Educational evaluation and decision making. Itasca, Ill.: Peacock.

Vallance, E. (1976) The application of aesthetic criticism to curriculum materials: Arguments and issues. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April, 1976.

Weiss, C.H. (1972) Evaluation research: Methods of assessing program effectiveness. Englewood Cliffs, N.J.: Prentice-Hall.

Wolf, R.L., Potter, J., and Baxter, B. (1976) The judicial approach to educational evaluation. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April, 1976.